Proceedings of the 18th Conference on Behavior Representation in Modeling and Simulation, Sundance, UT, 31 March - 2 April 2009
[paper: 09-BRIMS-018, 73-80]

The Synthetic Teammate Project

Jerry Ball[1], Christopher Myers[2], Andrea Heiberg[3], Nancy Cooke[2], Michael Matessa[4] & Mary Freiman[5]
Air Force Research Laboratory[1], Cognitive Engineering Research Institute[2]
General Motors[3], Alion[4], L3 Communications[5]
Jerry.Ball@mesa.afmc.af.mil, cmyers@cerici.org, ninaheib@cox.net, Nancy.Cooke@asu.edu,
mmatessa@alionscience.com, Mary.Freiman@mesa.afmc.af.mil

Keywords:
Synthetic Teammate, Language Comprehension/Generation, Dialog Manager, Situation Model, Task Behavior

ABSTRACT: *The main objective of the Synthetic Teammate project is to develop language and task enabled synthetic agents capable of being integrated into team training simulations. To achieve this goal without detriment in team training, the synthetic agents must be capable of closely matching human behavior. The initial application for the Synthetic Teammate research is the creation of an agent capable of performing the functions of a pilot for an Unmanned Aerial Vehicle (UAV) simulation as part of a three-person team.*

1. Project Overview

The main objective of the Synthetic Teammate project is to develop synthetic agents capable of being integrated into team training simulations. To achieve this goal without detriment in team training, the synthetic agents must be capable of closely matching human behavior across several cognitive capacities, such as situation assessment, task behavior, and language comprehension and generation. The initial application for the synthetic teammate research is the creation of an agent capable of functioning as the pilot of an Unmanned Aerial Vehicle (UAV) within a synthetic task environment (STE) which is described in the following section.

2. Synthetic Task Environment

The task environment used for developing the synthetic teammate is the Cognitive Engineering Research on Team Tasks (CERTT) UAV-STE (Cooke & Shope, 2005). The CERTT UAV-STE simulates teamwork aspects of UAV operations rather than equipment aspects (e.g., buttons and dials). The UAV-STE involves three interdependent team members, each with a different role. The team members are the Data Exploitation Mission Planning and Communications operator (DEMPC, the planning officer) who is responsible for creating a dynamic flight plan, including speed and altitude restrictions, an Air Vehicle Operator (AVO, the pilot) who controls flight settings and systems, and a Payload Operator (PLO, the sensor operator) who monitors sensor equipment and takes photographs.

The team members' common goal is to photograph ground targets and this requires interaction between all team members. Interaction occurs through a text-based communications system. A single UAV-STE mission consists of 11-12 ground targets and lasts a maximum of 40 minutes. However, a mission can end once the team photographs all possible targets.

The task requires a high degree of coordination due to time pressures and mutual constraints among the team member roles. To perform well within the UAV-STE, team members must understand their own tasks, and, more importantly, coordinate with each other to complete their common goal. The UAV-STE therefore provides an ideal task environment for developing a synthetic teammate.

3. Synthetic Teammate Overview

The Synthetic Teammate project is intended to lead to development of a cognitively plausible, yet functional synthetic teammate. The core of the system is being implemented within the ACT-R cognitive architecture (Anderson et al., 2004; Anderson, 2007), reflecting the focus on cognitive plausibility. As argued in Ball (2006), for inherently human behaviors like language comprehension (and generation), the use of a cognitive architecture to guide and constrain the implementation of a system may actually facilitate, rather than hinder, development. The constraints imposed by the cognitive architecture push system development in cognitively plausible directions which are more likely to lead to human-like behavior than purely algorithmic solutions which ignore such constraints. Although purely algorithmic solutions may provide short-term gains, they often lead to long-term difficulties as in a parser which processes the linguistic input from right to left—taking advantage of the punctuation at the end of a sentence—but can't be integrated with a speech recognition system or process language incrementally in real-time.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE
APR 2009

2. REPORT TYPE
Conference Proceedings

3. DATES COVERED
01-01-2008 to 30-03-2009

4. TITLE AND SUBTITLE
The Synthetic Teammate Project

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61101F

6. AUTHOR(S)
Jerry Ball; Christopher Myers; Andrea Heiberg; Nancy Cooke; Michael Matessa

5d. PROJECT NUMBER
2313

5e. TASK NUMBER
AS

5f. WORK UNIT NUMBER
09

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/RHA,Warfighter Readiness Research Division,6030 South Kent Street,Mesa,AZ,85212-6061

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL; AFRL/RHA

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL; AFRL/RHA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RH-AZ-PR-2009-0001

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Additional Author: Freiman, Mary Published in Proceedings of the 18th Conference on Behavior Representation in Modeling and Simulaiton, held in Sundance, UT, on 31 Mar-2 Apr 2009, Paper No. 09-BRIMS-018, 73-80

14. ABSTRACT
The main objective of the Synthetic Teammate project is to develop language and task enabled synthetic agents capable of being integrated into team training simulations. To achieve this goal without detriment in team training, the synthetic agents must be capable of closely matching human behavior. The initial application for the Synthetic Teammate research is the creation of an agent capable of performing the functions of a pilot for an Unmanned Aerial Vehicle (UAV) simulation as part of a three-person team.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Public Release**	**8**	

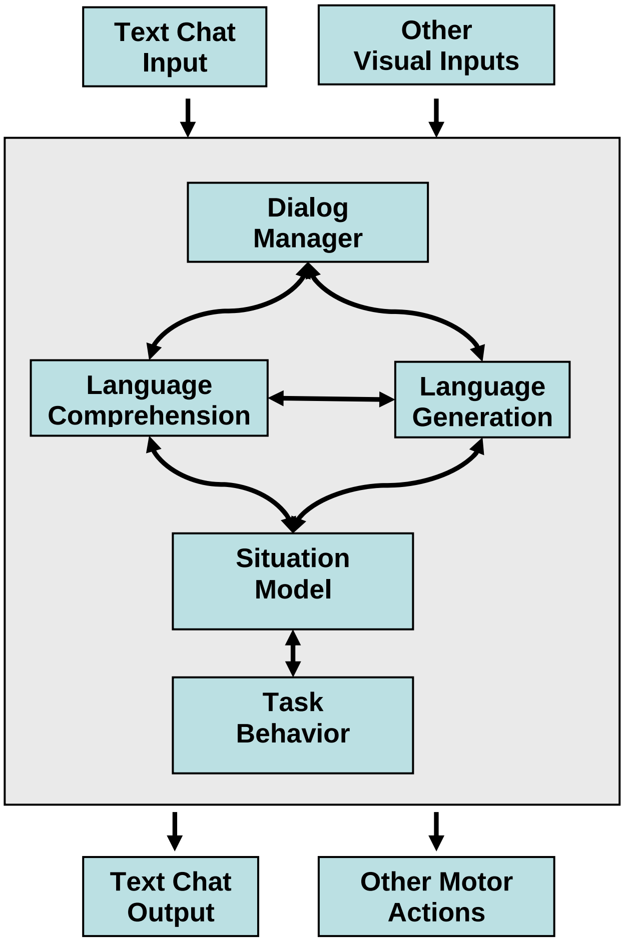

Figure 1: Synthetic Teammate Overview

The major linguistic components of the system include text chat based language comprehension and generation components, which are under the control of a dialog manager (see Figure 1). The linguistic subsystem interacts with a situation model component that is a spatial-imaginal/propositional representation of the current state of affairs as encoded from text chat inputs. The situation model component is intended to be a computational implementation of the notion of a situation model as described in Zwann & Radvansky (1998). The situation model component also reflects inputs from the visual system via the task behavior component. The task behavior component implements the behavior of the system, controlling shifts of attention in the visual system and motor actions needed to perform the pilot's tasks. Input to the system is mediated by ACT-R's perceptual module and motor actions are mediated by ACT-R's motor module. The perceptual and motor modules are ACT-R's interfaces to the external environment. Each of the model components makes use of ACT-R's declarative memory and production system.

Most of the current research has been focused on individual development of the language comprehension component, language generation & dialog manager components, and task behavior component. The language generation & dialog manager components, which were developed jointly, have recently been integrated with the task behavior component via a "situation superchunk" which contains the knowledge needed and generated by the components. The situation superchunk will eventually be replaced by the situation model component, currently being designed. The following sections provide more detail for each of the synthetic teammate's core components.

4. Language Comprehension Component

The language comprehension component has been under development since the mid 1980's (Ball, 1991) with a hiatus in the 90's. Originally developed in Prolog, the language comprehension component was ported to the ACT-R 5 architecture in 2003 (Ball, 2004a). The current version runs in ACT-R 6 (Ball, Heiberg, & Silber, 2007). The language comprehension component is intended to be a domain general system capable of handling a wide range of English constructions. There is no assumption that the specific domain of application can be used to limit the scope of the system. Additions to the model to handle the text-chat specific corpus are being made in the context of a regression testing capability to insure that the broad coverage of the component is maintained.

The language comprehension component is a construction-driven processing system (Ball, 2007a) based on a linguistic theory of the grammatical encoding of referential and relational meaning (Ball, 2007b). The linguistic theory is aligned with basic principles of Cognitive and Construction Grammar (cf. Langacker, 1987, 1991). Lexical items in the linguistic input activate constructions which drive processing. For example, the transitive verb "increase" activates a transitive verb construction. This construction, if selected, sets up an expectation for an object to occur. The transitive verb construction also projects a clausal construction (if one hasn't already been projected by a preceding auxiliary verb). The clausal construction sets up an expectation for a subject. The subject of the clausal construction is typically available in the current context and, if available, is integrated into the clausal construction. The absence of a subject can trigger projection of an imperative clause construction if the verb is in the base form as in "increase the altitude", otherwise a declarative clause construction is projected even if the subject is missing (e.g., "increased the altitude"). The occurrence of an auxiliary verb preceding the subject can trigger projection of a yes-no question construction as in “are you increasing the altitude”. If a wh-word precedes the verb, a wh-question construction is projected as in “who increased the altitude” or “why are you increasing the altitude”.

The language comprehension component processes the input incrementally (one word at time), constructing a linguistic representation of the input based on the current word, constructions activated by the word, and the prior context. If necessary, the current input is accommodated by adjusting the current representation or coercing the current input into that representation without backtracking or lookahead. The mechanism of *context accommodation* is part and parcel of the basic left-to-right, incremental processing mechanism. For example, in the processing of “the airspeed restriction”, when “airspeed” is processed it is integrated as the head of the nominal construction projected by “the”. However, when the word “restriction” is processed, the nominal construction is adjusted so that “airspeed” functions as a modifier, with “restriction” functioning as the head. Context accommodation avoids the need to carry forward multiple representations in parallel, and yet the model still arrives at an appropriate representation at the end of processing.

The language processor is highly context sensitive and makes use of all available information—lexical, syntactic, semantic and pragmatic—in deciding how to process a given input. There is no autonomous syntactic component or syntactic processor, although grammatical information is very important for determining meaning. Contextual information is probabilistically summed via ACT-R’s parallel spreading activation mechanism to yield the best alternative given the current input and context. The selected alternative is assumed to be correct and the processor proceeds deterministically and serially forward. The context sensitive, probabilistic, parallel, spreading activation mechanism, combined with a mechanism of context accommodation makes a nearly deterministic, serial language processing system possible.

Recent modifications to the language comprehension component have focused on the processing of long-distance dependencies—demonstrating that the system is capable of handling such theoretically important constructions (e.g. the theoretically important examples “he$_i$ is eager t$_i$ to please” vs. “he$_i$ is easy to please t$_i$”).

The language comprehension component is also being extended to handle the text-based communication corpus that was collected in an experiment involving human subjects and the UAV-STE. The text chat corpus is full of interesting variability in the form of linguistic input (e.g. typos, spelling variants, morphological variants, abbreviations, acronyms, concatenations, new coinages). In order to handle this variability, lower level processes of word recognition have been added to the language comprehension component. The spreading activation mechanism of ACT-R allows the model to retrieve words from the lexicon that are not an exact match to the input. Letters and trigrams in the input spread activation to the words containing those letters and trigrams in the mental lexicon. These processes and encodings are based on the Interactive Activation model of word recognition (McClelland and Rumelhart 1981), with the addition of trigrams based on the “letter triples” as later described by Seidenberg and McClelland (1989). Though inspired by the findings of word recognition studies, this subcomponent of the model does not model a word recognition task. It is embedded in the language comprehension component as a whole; therefore, the effects of context and previous activation levels must be taken into consideration when encoding each individual word (Freiman & Ball, in press).

5. Language Generation & Dialog Manager Components

The language generation and dialog manager components were developed to capture the dynamic nature of human language production, following earlier approaches involving dynamic dialog constraints (Ericcson, 2004), accommodation (Matessa, 2000), and adaptive content selection (Walker et al., 2004). The focus is on selecting the best utterance from a set of possible utterances which were derived from a UAV-STE experiment involving spoken communication. The approach is akin to overgeneration-and-ranking approaches (Varges, 2006).

The model uses Optimality Theory (Prince & Smolensky, 1993; 2004) to select an optimal utterance, given a set of utterances and a set of constraints on utterances. Constraints are simple, violable, conflicting, and motivated by cross-linguistic evidence. Constraints are arranged in a strict dominance hierarchy; the optimal utterance is the one that least violates the hierarchy.

Constraint ranking is expressed through ACT-R spreading activation. Activation spreads from constraints to utterances to determine which utterance is retrieved from memory. The most important constraint spreads the most activation and has the greatest effect on retrieval. Factors from the situation component dynamically affect the constraint ranking, possibly reranking constraints, and providing a principled variation in utterances over time.

The language generation component is based on retrieval of complete utterances with one or two variabilized slots. These utterance templates are akin to constructions, but there is currently no capability to integrate multiple constructions together, as in the language comprehension component. Purely constraint based approaches like OT are good at selecting among competing alternatives, but require additional mechanisms to support productive generation of alternatives from smaller linguistic units.

The dialog manager component models the push and pull of information to and from the AVO. It uses a temporal module extension to ACT-R to avoid repeatedly asking for the same information.

6. Task Behavior Component

The task behavior component was developed to fly the UAV from waypoint to waypoint in a cognitively plausible manner. Flying to waypoints involves interacting with the UAV-STE to queue the correct waypoint and enter the correct course. The pilot must also set the UAV airspeed and altitude within restrictions provided by the sensor operator (PLO) and planning officer (DEMPC). The task model interacts with the UAV-STE using the same devices as humans—it uses the mouse pointer to interact with the UAV flight controls in a point-and-click fashion, and uses the keyboard to send and receive messages to and from its teammates.

The task model was developed using a combination of hierarchical task analysis and NGOMSL (Kieras, 1988). The analysis identified the goals necessary for accomplishing flight from one waypoint to another, the sequence flexibility of the goals, and commonalities across all goals.

The goals associated with the task behavior component include setting flight parameters (i.e., altitude, speed, and course), setting waypoints, monitoring alarms and warnings, and monitoring the UAV flight status (i.e., the distance from upcoming waypoint and the time to the next waypoint, etc.). Each of these goals is divided into three subgoals, *obtaining desired state information, checking current state information* and *changing the current state to the desired state*. Each subgoal updates the appropriate information within the situation component (i.e., situation superchunk).

The first component, *obtaining*, is modeled to obtain the desired state information. Once this is done, the second component, *checking*, is executed to determine if the desired state differed from the current state. When there is a discrepancy, the model performs the third component, *changing*, to modify the task to a desired state. As a result of breaking each of the task goals into three components, there has been a substantial re-use of production rules within the task model.

For example, assume the task behavior component has received the next waypoint from the planning officer. This information is stored in the situation model component and used to retrieve the goal from memory for checking waypoint information. To check the next waypoint value, the model attends and encodes the "queued waypoint" value on the GUI and determines if the queued waypoint needs to be adjusted (i.e., obtaining and checking). If the waypoint needs to be adjusted, then the task model spawns a goal to attend to the waypoint setting information and set the desired waypoint using the appropriate mechanism (i.e., changing).

7. Situation Model Component

The Situation Model component represents the current situation as informed by the linguistic input, the task environment, the discourse context, and salient background knowledge. The situation model constitutes the primary meaning representation of the system, although the linguistic representations that get mapped into the situation model also encode important aspects of meaning. The situation model component is responsible for grounding the meaning of referring expressions in the linguistic input in the objects and situations from the task environment, discourse context and background knowledge which are encoded in the situation model.

The concept of a *Situation Model* originates in the research of Kintsch and van Dijk (1978) and corresponds to a mental representation of the propositional content of a text—including the addition of propositions corresponding to inferences that are derived from the text. The term "situation model" implies that this propositional representation is a model of the situation described in the text. For example, given the text "he put the book on the table" a propositional representation like

PUT(JOHN,ON(BOOK,TABLE))

(where "he" is resolved to refer to John and the use of uppercase words correspond to concepts) might be generated. Note that this representation contains the inference that the book is on the table. The mapping from a linguistic text to a propositional representation of the corresponding situation has not been fully automated in the computational research of Kintsch (cf. Kintsch, 1998). Later psychological research on situation models has established that the mental representation of situations corresponding to texts contain spatial-imaginal and temporal information, as well as propositional information (cf. Zwann & Radvansky, 1998). However, there are no computational accounts of how spatial-imaginal information is represented in a situation model.

We are currently in the process of developing an initial design for the situation model and the discussion in this section is preliminary and subject to change. However, a considerable amount of time, effort and resources have already been committed to this project and despite the preliminary nature of this system component, this project is well advanced by any reasonable measure for complex system development.

7.1 Propositional Content

In terms of representing propositional content, we adhere to the principle that the propositional (or logical) notation should be as close to English as possible (Hobbs, 1985). In this regard, the predicates used in the propositional representations are concepts that correspond to English words and are referred to as "word-concepts". The primary distinction between a word and a word-concept is not based on the idea that concepts are non-linguistic or pre-linguistic, but that words are organized into an ontology which reflects their grammatical function, whereas word-concepts are organized into an ontology with reflects their semantic content.

In this regard, we are considering the use of WordNet synonym sets (cf. Miller, 1995) as the source of word-concepts. For example, the word "raise" is grammatically categorized as a transitive verb, whereas the word-concept "raise-1-cncp" is semantically categorized as a change verb and "raise-2-cncp" is categorized as a motion verb in WordNet—in two common verb senses of "raise". The word "raise" participates in linguistic processing and the generation of linguistic representations, whereas the word-concepts "raise-1-cncp" and "raise-2-cncp" participate in situation model processing and in the generation of situation model representations. In the simplest case, there is a direct mapping from word to word-concept and the generation of a situation model representation from a linguistic representation is facilitated. However, besides often having multiple senses that need to be disambiguated to do the mapping, it may be that words map into word-concepts based on a synonym of a word, rather than the word itself. For example, the word "radius" as used in "the effective radius is 5 miles"—which indicates the region around a waypoint at which a picture may be taken—may map into a "region-cncp" which could be used as the word-concept label for the WordNet synonym set for this sense of "radius". The alternative of using WordNet synset id's like 08628578 to represent this sense of "radius" is unattractive from a representational perspective. Another possibility is to tag the word with the synset id as in "radius-08628578-cncp". In this case, "region" could also be tagged with the same id "region-08628578-cncp" to indicate their synonymy.

Besides specifying the nature of word-concepts corresponding to predicates, we need to specify how these predicates are integrated together into complex representations, and, ultimately, how these representations are mapped into the representational formalism of the ACT-R architecture which is essentially frame based—i.e. declarative memory (DM) chunks are named and typed sequences of slot-value pairs organized into a single inheritance hierarchy. We plan to borrow ideas from Hobbs (1985, internet) and Discourse Representation Theory (Kamp & Ryle, 1993) in the design of our propositional system of representation. In terms of the mapping to ACT-R DM chunks, an initial attempt to specify a mapping from the Cyc ontology of concepts into ACT-R declarative memory chunks is described in Ball, Rodgers & Gluck (2004). An outcome of that research was the realization that the Cyc ontology does not provide the domain specific concepts needed in our particular task domain. Many of the domain specific concepts have now been identified via analysis of the text chat corpus and task domain.

7.2 Spatial Content

To represent spatial aspects of the situation, we plan to use a spatial module developed for use with ACT-R and described in Douglass (2007). This module is designed to support the mental representation of objects and spatial relations between objects in a graphical display. An obvious use of this module is for spatially representing the graphical objects in the three monitors that constitute the graphical user interface (GUI) of the AVO. Another possible use is to represent the sequence of waypoints that must be visited during a reconnaissance mission.

7.3 Imaginal Content

There is abundant evidence that humans reason over imaginal representations (cf. Kosslyn, 2006; Zwann & Radvansky, 1998) and our task domain strongly suggests the need for such a capability. However, a computational implementation of an imaginal reasoning capability is currently outside the scope of the project—even though eventual development of such a capability is important for attaining full cognitive plausibility.

7.4 Discourse Content

A representation of the discourse participants (e.g. PLO, DEMPC, Intel Officer, AVO) is crucial to development of a functional synthetic teammate, as is a capability to determine the discourse acts that are inferable from the linguistic inputs. For example, when the PLO sends the message “I need to be above 3000” to the AVO, the AVO must infer that this is a request to increase the altitude of the UAV to be above 3000 feet, despite the fact that the linguistic input is a declarative statement which is ostensibly about the PLO, not the UAV, and there is no mention of what “3000” quantifies.

As the discourse advances across missions, human teammates adapt to each other’s communications, standardizing forms and providing less and less explicit content in the messages. An adaptive capability to adopt standard forms and to infer implicit information from the evolving discourse context is needed (Matessa, 2000). That adaptive capability will hinge on the information available in the situation model. We would also like the synthetic teammate to be capable of reasoning about the mental state of the other team members, but this is currently outside the scope of our development efforts.

8. Scaling up the Cognitive Architecture

ACT-R was designed to support the development of small-scale cognitive models of specific laboratory phenomena. Since the advent of the first computational version of ACT-R, hundreds of small-scale models have been developed. The synthetic teammate project is one of a few attempts to develop a larger-scale model (or system of models) in ACT-R. This development is pushing ACT-R in directions for which it was not originally designed. For example, the parallel spreading activation mechanism of ACT-R is computationally explosive on serial hardware. To support the computation of the activation of DM chunks corresponding to thousands of lexical items, we have integrated a relational database with ACT-R. The relational database allows us to externalize ACT-R’s DM and provides highly efficient database retrieval mechanisms that are allowing us to expand the model’s mental lexicon to a reasonable size. Further, the integration of a relational database allows us to maintain declarative knowledge acquired over many model runs—a capability not previously available in ACT-R.

The current language comprehension component contains over 2500 words in its mental lexicon. We plan to increase this substantially via integration of additional words from the WordNet mental lexicon which contains > 100,000 words. For this project, we expect to need 10-15,000 words in the mental lexicon. Efforts are currently underway to map the entries in WordNet into the form needed by the language comprehension model. The mapping of nouns, adjective and adverbs is straightforward and can be automated, but the mapping of verbs with their varying argument structures is more problematic. Currently the model has some capability for word sense disambiguation (WSD), but the addition of a full-size mental lexicon will stress this capability beyond its limits. We are evaluating the use of Latent Semantic Analysis (cf. Landauer & Dumais, 1998) to provide additional WSD capability. In addition, it is not enough to just have a large lexicon. The model must be capable of taking appropriate action giving the linguistic input, and this requires a deeper level of understanding than is typical of most wide coverage, but superficial, computational linguistic systems.

9. Empirical Validation

An important goal of the project is to develop a synthetic teammate that is at once functional and cognitively plausible. In a system as complex as the synthetic teammate, empirical validation is a significant challenge. It is not practical to individually validate all the possible behaviors of the system. Instead, a few key behaviors will be selected for scrutiny and validated against empirical data. At the highest level, we will determine whether or not teams with a synthetic AVO show evidence of learning that all human teams in the UAV-STE demonstrate. We also plan to compare the communicative behavior of the synthetic teammate in terms of the “push” and “pull” of information against data that has been collected for human teams. It should be noted that this empirical validation will occur within the context of a functioning synthetic teammate, an atypical empirical approach which will lend credibility to the model in the sense that the model must do much more than just show evidence for aligning with a specific data set – the model must also function as a teammate with all the constraints on model behavior which that entails.

Furthermore, it is an empirical goal of the language comprehension component to be able to process linguistic input in real-time on Marr’s algorithmic level (Marr, 1982) where parallel and serial processing mechanisms are relevant (Ball, 2008). This goal imposes serious constraints on possible processing mechanisms—for example, eliminating non-deterministic mechanisms that rely on algorithmic backtracking and cannot, in principle, operate in real-time since such mechanisms slow down with the length of the linguistic input.

Finally, not all components of the synthetic teammate are equally cognitively plausible. In the interest of building an end-to-end system, cognitive constraints on the development of the language generation and dialog manager components have been relaxed. Although less cognitively plausible, these components do a good job of modeling the language generation behavior of the individual AVO on which they were modeled. On the other hand, the task behavior component, which takes advantage of the perceptual-motor modules of ACT-R, is more closely tied to cognitive plausibility—down to the timing of attention fixations, key presses and mouse movements.

10. Comparison to Other Approaches

The use of the term “Synthetic Teammate” is borrowed from research ongoing at Chi Systems (cf. Scolaro & Santarelli, 2002). In a panel session at BRIMS in 2004, there were presentations of several different approaches to the development of synthetic agents with natural language capabilities (Ball, 2004b). The Synthetic Teammate project aligns with this research. However, unlike other systems, the Synthetic Teammate project is based on text chat rather than spoken input. The challenges of processing spoken language limit the capabilities of spoken language systems (Stokes, 2001). Such systems typically assume a restricted vocabulary and limited forms of input in order to cope with this challenge. We have decided to use text chat to overcome these limitations. A similar approach has been adopted in the Situation Understanding BOT thru Language and Environment (SUBTLE) project (Marcus, et al., 2008). However, the SUBTLE project has the additional challenge of having to situate the synthetic teammate on a robot platform and act in the real world.

The defining feature of this research is the focus on cognitive plausibility, often at a fine-grained level of cognitive fidelity uncharacteristic of most research in the development of synthetic agents.

11. Conclusions

The Synthetic Teammate project is a challenging project reminiscent of earlier research in Artificial Intelligence and Cognitive Science which focused on solving AI Hard Problems using cognitively motivated computational techniques. The current goal is to have an initial end-to-end system in place by summer 2009. The initial system will be subjected to iterative refinement until a version which is capable of functioning as a teammate in the UAV-STE simulation is available. Once reasonable functionality is achieved, an experiment will be conducted in which the synthetic teammate will interact with human teammates, and the performance of this hybrid team will be compared against all human teams at the individual and team levels of analysis.

11. Acknowledgements

This research is funded in part by applied research funds from the Warfighter Readiness Research Division, Human Effectiveness Directorate, 711th Human Performance Wing, Air Force Research Laboratory and by basic research funds from the Air Force Office of Scientific Research and Office of Naval Research. The Synthetic Teammate project is a sizable project involving collaboration between the Air Force Research Laboratory and the Cognitive Engineering Research Institute. Project team members not listed as authors on this paper—but who have contributed to the project in important ways—include Dee Andrews, Scott Douglass, Jasmine Duran, Kevin Gluck, Jack Harris, Mike Krusmark, Don Lyon, Harry Pedersen, Eric Robinson, Stuart Rodgers, Steven Shope, Ronnie Silber and Amanda Taylor.

12. References

Anderson, J. R. (2007). *How Can the Human Mind occur in the Physical Universe?* NY: Oxford.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S, Lebiere, C, and Qin, Y. (2004). An Integrated Theory of the Mind. *Psychological Review* 111, (4). 1036-1060

Ball, J. (1991). PM, Propositional Model, a Computational Psycholinguistic Model of Language Comprehension Based on a Relational Analysis of Written English, Ann Arbor, MI: UMI Dissertation Information Service.

Ball, J. (2004a). A Cognitively Plausible Model of Language Comprehension. *Proceedings of the 13th Conference on Behavior Representation in Modeling and Simulation*, pp. 305-316. ISBN: 1-930638-35-3

Ball, J. (2004b). Symposium: Software Agents with Natural Language Capabilities–Where are we? *Proceedings of the 13th Conference on Behavior Representation in Modeling and Simulation*, xxix-xxxi. ISBN: 1-930638-35-3

Ball, J. (2006). Can NLP Systems be a Cognitive Black Box? *Papers from the AAAI Spring Symposium, Between a Rock and a Hard Place, Cognitive Science Principles Meet AI Hard Problems.* Technical Report SS-06-02: AAAI Press.

Ball, J. (2007a). Construction-Driven Language Processing. *Proceedings of the 2nd European Cognitive Science Conference*, pp. 722-727. Edited by S. Vosniadou, D. Kayser & A. Protopapas. NY: LEA

Ball, J. (2007b). A Bi-Polar Theory of Nominal and Clause Structure and Function. *Annual Review of Cognitive Linguistics*, 27-54. Amsterdam: John Benjamins

Ball, J. (2008). A Naturalistic, Functional Approach to Modeling Language Comprehension. *Papers from the AAAI Fall Symposium, Naturally Inspired Artificial Intelligence.* Technical Report FS-08-06: AAAI Press.

Ball, J., Heiberg, A. & Silber, R. (2007). Toward a Large-Scale Model of Language Comprehension in ACT-R 6. *Proceedings of the 8th International Conference on Cognitive Modeling*, pp. 163-168. Edited by R. Lewis, T. Polk & J. Laird.

Ball, J., Rodgers, S., & Gluck, K. (2004). Integrating ACT-R and Cyc in a Large-Scale Model of Language Comprehension for Use in Intelligent Agents. *Papers from the AAAI Workshop, Technical Report WS-04-07*, pp. 19-25. Menlo Park, CA: AAAI Press

Cooke, N. & Shope, S. (2005). Synthetic Task Environments for Teams: CERTT's UAV-STE. *Handbook on Human Factors and Ergonomics Methods*. 46-1-46-6. Boca Raton, FL: CLC Press, LLC

Douglass, S. (2007). A Computational Model of Situated Action. Dissertation. Carnegie Mellon University.

Freiman, M., & Ball, J. (in press). Computational cognitive modeling of reading comprehension at the word level. *Proceedings of the 38th Western Conference on Linguistics*.

Ericsson, S. (2004). Dynamic optimisation of information enrichment in dialogue. In *Proceedings of the 8th International Workshop on Formal Semantics and Pragmatics of Dialogue*, Catalog, Barcelona, Spain.

Hobbs, J. (1985). Ontological Promiscuity. *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, 61-69.

Hobbs, J. (internet). Discourse and Inference. Downloaded Jan 2009 from http://www.isi.edu/~hobbs/disinf-tc.html

Kamp, H. & Ryle, U. (1993). *From Discourse to Logic*. Dordrecht: Kluwer.

Kieras, D. E. (1988). Towards a practical GOMS model methodology for user interface design. In M. Helander (ed.), *The handbook of human-computer interaction*, 135-158.Amsterdam: North-Holland.

Kintsch, W. (1998). *Comprehension, a Paradigm for Cognition*. NY: Cambridge University Press

Kintsch, W. & van Dijk, T. (1978) Towards a model of text comprehension and production. *Psychological Review, 85*, 363-94.

Kosslyn, S. (2006). *The Case for Mental Imagery*. NY: Oxford University Press.

Landuaer, T. & Dumais, S. (1997). A Solution to Plato's Problem: The Latent Semantic Analysis Theory of the Acquisition, Induction, and Representation of Knowledge. Psychological Review, 104:2, 211-240.

Langacker, R. (1987, 1991). *Foundations of Cognitive Grammar*, Volumes 1 & 2. Stanford, CA: Stanford University Press.

Marcus, M. et al. (2008). SUBTLE project program review.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: Freeman.

Matessa, M. (2000). Simulating Adaptive Communication. Carnegie Mellon University Ph.D. dissertation.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part I. An account of basic findings. *Psychological Review,* 88, 375-407.

Miller, G. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38, 11, 39-41.

Myers, C. W., Cooke, N. J., Ball, J. T., Heiberg, A., Gluck, K. A., & Robinson, F. E. (under review). Collaborating with Synthetic Teammates. In W. Bennett (ed.), *Collaboration in Complex Task Environments*.

Prince, A. & Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell.

Scolaro, J. & Santarelli, T. (2002). Cognitive modeling teamwork, taskwork, and instructional behavior in synthetic teammates. *Proceedings of the 11th Conference on Computer Generated Forces and Behavioral Representation*. Orlando: Institute for Simulation and Training.

Seidenberg, Mark S., & McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review*, 96, No. 4, 523-568.

Stokes, J. (2001). Speech interaction and human behavior representations (HBRs). *Proceedings of 10th Conference on Computer Generated Forces and Behavioral Representation*, 467-476. Norfolk, VA: SISO, Inc.

Varges, S. (2006). Overgeneration and ranking for spoken dialogue systems. *Proceedings of the 4th International Natural Language Generation Conference*, 20–22, Sydney, Australia, July 2006. Association for Computational Linguistics.

Walker, M. A., S. J. Whittaker, A. Stent, P. Maloor, J. Moore, M. Johnston, & G. Vasireddy. (2004). Generation and evaluation of user tailored responses in multimodal dialogue. *Cognitive Science* 28:5, 811-840.

Zwann, R., and Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162-185.

Author Biography

Dr. Jerry Ball is a senior research psychologist in the Human Effectiveness Directorate, 711th Human Performance Wing, Air Force Research Laboratory

Dr. Christopher Myers is a post-doc at the Cognitive Engineering Research Institute (CERI).

Dr. Andrea Heiberg was previously a research scientist at L3 Communications.

Dr. Nancy Cooke is the lead scientist at CERI and a professor at Arizona State University – Polytechnic.

Dr. Michael Matessa is a research scientist at Alion.

Mary Freiman is a research scientist at L3 Communications.